Biodemography and Social Biology, 58:14–39, 2012

ISSN: 1948-5565 print / 1948-5573 online
DOI: 10.1080/19485565.2012.666121

Routledge
Taylor & Francis Group

Biodemography of Exceptional Longevity: Early-Life and Mid-Life Predictors of Human Longevity

LEONID A. GAVRILOV AND NATALIA S. GAVRILOVA

Center on Aging, NORC at the University of Chicago, Chicago, Illinois, USA

This study explores the effects of early-life and middle-life conditions on exceptional longevity using two matched case-control studies. The first study compares 198 validated centenarians born in the United States between 1890 and 1893 to their shorter-lived siblings. Family histories of centenarians were reconstructed and exceptional longevity validated using early U.S. censuses, the Social Security Administration Death Master File, state death indexes, online genealogies, and other supplementary data resources. Siblings born to young mothers (aged less than 25 years) had significantly higher chances of living to 100 compared to siblings born to older mothers (odds ratio = 2.03, 95% CI = 1.33–3.11, p = .001*). Paternal age and birth order were not associated with exceptional longevity. The second study explores whether people living to 100 years and beyond differ in physical characteristics at a young age from their shorter-lived peers. A random representative sample of 240 men who were born in 1887 and survived to age 100 was selected from the U.S. Social Security Administration database and linked to U.S. World War I civil draft registration cards collected in 1917 when these men were 30 years old. These validated centenarians were then compared to randomly selected controls who were matched by calendar year of birth, race, and place of draft registration in 1917. Results showed a negative association between "stout" body build (being in the heaviest 15 percent of the population) and survival to age 100. Having the occupation of "farmer" and a large number of children (4 or more) at age 30 increased the chances of exceptional longevity. The results of both studies demonstrate that matched case-control design is a useful approach in exploring effects of early-life conditions and middle-life characteristics on exceptional longevity.*

Introduction

Studies of centenarians (persons living to age 100 and over) are useful in identifying factors leading to long life and avoidance of fatal diseases. Even if some childhood or middle-life factors have a moderate protective effect on the risk of death, this longevity advantage

This study was supported by the U.S. National Institute on Aging (Grant R01 AG028620). We are grateful to two anonymous reviewers for their constructive criticism and useful suggestions. Earlier results of this study were presented and discussed at the 2006 and 2008 annual meetings of the Population Association of America, the 2008 annual meeting of the Gerontological Society of America, the 2008 Society of Actuaries "Living to 100 and Beyond" international symposium, and the 2007 and 2009 annual meetings of the International Network on Health Expectancy and the Disability Process (REVES).

Address correspondence to Leonid A. Gavrilov, Center on Aging, NORC at the University of Chicago, 1155 East 60th Street, Chicago, IL 60637, USA. E-mail: gavrilov@longevity-science.org

would result in the accumulation of long-lived persons who experienced the effects of these factors. Thus, the study of centenarians may be a sensitive way to find genetic, familial, environmental, and life-course factors associated with lower mortality and better survival.

Most studies of centenarians in the United States focus on either genetic (Barzilai and Shuldiner 2001; Barzilai et al. 2006; Hadley et al. 2000; Puca et al. 2001) or psychological (Adkins, Martin, and Poon 1996; Hagberg et al. 2001; Jang, Poon, and Martin 2004; Martin, da Rosa, and Siegler 2006) aspects of survival to advanced ages. However, several theories suggest that early-life events and conditions may have significant long-lasting effects on survival to advanced ages (Barker 1992; Fogel and Costa 1997; Gavrilov and Gavrilova 1991; Kuh and Ben-Shlomo 1997). These ideas are supported by a number of studies that demonstrate the significant effects of early-life conditions on late-life mortality (Alter and Oris 2005; Ben-Shlomo and Kuh 2002; Bengtsson and Mineau 2009; Costa 2003; Elo and Preston 1992; Gavrilov and Gavrilova 2003; Hayward and Gorman 2004; Smith, Mineau et al. 2009). The existence of correlations between early growth patterns and subsequent fitness has been established for both human beings and several other mammalian species (Lummaa and Clutton-Brock 2002). Centenarian studies may be a useful way to identify early-life factors and conditions that affect survival to advanced ages.

At the same time, studies of centenarians face significant difficulties in collecting reliable data and finding appropriate study design and methodology. Survival to age 100 is a rare event (only two men and 14 women out of one thousand from a 1900 birth cohort survived to age 100), and therefore traditional methods of population-based sampling are not feasible for obtaining large samples of centenarians. The case-control design, however, has proved to be the most appropriate and cost-effective approach for studies of rare conditions (Breslow and Day 1993; Woodward 2005) and is thus extremely useful for centenarian studies. The classic case-control design can be expanded in a variety of ways, with one such expansion being a design suggested by Preston, Hill, and Drevenstedt (1998). In this design, the case is identified as survival to advanced ages (rather than disease or death), and relative survival probabilities are used instead of odds ratios.

Studies of the effects of early-life conditions on exceptional longevity face additional difficulties from possible confounding resulting from between-family variation in childhood socioeconomic conditions and parental genetic background. It has been suggested that one possible solution to these challenges is to compare associations within sibships while adjusting for the fact that socioeconomic and genetic backgrounds are similar for siblings from the same family (Gavrilov and Gavrilova 2001; Smith, Gagnon et al. 2009).

In this article, we present two studies that use matched case-control design to analyze the effects of early-life and middle-life characteristics on individuals' survival to age 100. In the first study, we attempt to find out why centenarians differ from their shorter-lived siblings. In the second study, we compare centenarian characteristics recorded at a young age to similar characteristics of peers who did not survive to this advanced age. For both studies, we describe new historical data resources that proved useful in conducting our research, as well as the potential problems of handling these data.

Study 1: Why Centenarians Are Different from Their Shorter-Lived Siblings

The within-family design may be useful for studying the effect of parental age at reproduction (and other characteristics that vary within the same family) on exceptional longevity. While the detrimental effects of late reproduction on adverse reproductive outcomes and

genetic diseases has been well documented (Bottini et al. 2001; Gavrilov and Gavrilova 1997; Pellestor, Anahory, and Hamamah 2005), less is known about the long-term effects of delayed parenting on the health and longevity of adult offspring. Empirical evidence shows that the quality of female eggs in human beings rapidly declines with age (Bickel 2005; Pellestor, Anahory, and Hamamah 2005), and that this deterioration starts rather early, before age 30 (Heffner 2004). However, our previous studies did not detect an association between maternal age and offspring mortality in historical populations of European aristocracy (Gavrilov and Gavrilova 1997; Gavrilov and Gavrilova 2000; Gavrilov and Gavrilova 2001). Crow (1997; 2000) has suggested that paternal age at reproduction is the main factor determining human spontaneous mutation rate. A number of studies have demonstrated that advanced paternal age may be detrimental to the survival of adult offspring (Gavrilov and Gavrilova 1997; Gavrilov and Gavrilova 2000; Smith, Mineau, et al. 2009), although other studies have failed to show this effect (Hubbard, Andrew, and Rockwood 2009; Robine et al. 2003).

We believe that the within-family design is a useful way to study parental age effects on offspring longevity. This article presents the results of applying a within-family design to the study of the effects of parental age and birth order on exceptional longevity. For this purpose, we collected family histories of 671 putative centenarians born in the United States between 1890 and 1899 and conducted thorough validation of their birth and death dates. Then, for a subsample of 198 centenarians born between 1890 and 1893, we carried out an additional search for detailed information about all sample members' siblings. This procedure allowed us to create a sample of centenarians with information on their siblings' lifespans.

Study Design

This study explored the effects of early-life factors (birth order, paternal age, maternal age) on individuals' likelihood of survival to an advanced age. Centenarians (cases) were compared to their "normal" shorter-lived siblings (controls) using a within-family approach.

The study applied a case-sibling design (see Figure 1), a variant of a matched case-control design in which siblings of cases (long-lived individuals) are used as controls (Woodward 2005). This approach allows investigators to study within-family differences without being confounded by between-family variation. Long-lived persons born between 1890 and 1893 were used as cases. Siblings born no more than 10 years apart from the cases were used as controls to maximize similarity in family conditions and minimize the risk of lifespan being unknown for later-born siblings who were still alive by the time of data collection and hence could survive to age 100 in the future. As a result, controls were born between 1881 and 1902, and the sample included only three persons born after 1899 for whom we had no information on death dates and who could potentially be alive at the time of our study and survive to age 100. Taking into account the rarity of survival to age 108, it is safe to assume that these persons were dead by the time of our study and did not become centenarians. Therefore, we believe that centenarian status was correctly established for almost all siblings in our study.

The main approach used in this study is based on comparison of children within, rather than across, families. Within a family, children are born to parents at different ages, and this variation may be used to estimate the net effect of parental age more conclusively (Kalmijn and Kraaykamp 2005).

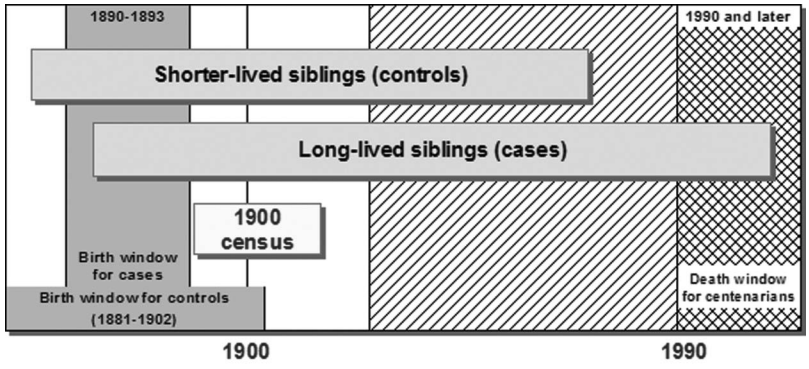

Figure 1. Description of case-sibling design.
Note: Circles indicate births and diamonds indicate deaths.

Data Collection and Quality Control

Family histories (genealogies) have long been a useful source of information for studies in historical demography (Adams and Kasakoff 1984; Adams and Kasakoff 1991; Anderton et al. 1984; Anderton et al. 1987; Bean, Mineau, and Anderton 1992; Kasakoff and Adams 2000) and biodemography (Gavrilov and Gavrilova 2001; Gavrilov et al. 2002; Kerber et al. 2001). In study, data on long-lived individuals (persons surviving to age 100 and beyond) were collected from computerized family histories available online at Rootsweb.com. Specifically, we extracted family data for 671 alleged centenarians born in the United States between 1890 and 1899 from publicly available computerized family histories of about 75 million individuals identified earlier by Gavrilova and Gavrilov (1999).

Data quality control is an important part of all centenarian studies, and in our case, it included (1) preliminary quality control of computerized family histories (data consistency checks), (2) verification of centenarians' death date, and (3) verification of birth dates for centenarians and their siblings (controls). All records for both centenarians and controls were subjected to verification and quality control procedures using independent data sources. We were primarily concerned with the possibility of incorrect reporting of dates in family histories. Previous studies have demonstrated that age misreporting and exaggeration in particular are more common among long-lived individuals (Elo et al. 1996; Hill et al. 2000; Jeune and Vaupel 1999; Rosenwaike and Hill 1996; Rosenwaike and Stone 2003; Rosenwaike et al. 1998; Shrestha and Rosenwaike 1996; Young et al. 2010). Therefore, we were careful to verify the ages of long-lived individuals in our study. To do so, we followed the age verification and data linkage approach developed by a team of demographers at the University of Pennsylvania (Elo et al. 1996; Hill et al. 2000; Preston et al. 1996; Rosenwaike and Hill 1996; Rosenwaike and Stone 2003; Rosenwaike et al. 1998). This approach uses the Social Security Administration Death Master File (DMF) to verify death dates and early U.S. censuses to verify birth dates. In this study, records from family histories were first linked to DMF records for death date validation and then to U.S. census records for the years 1900, 1910, and 1920 for birth date validation.

The DMF is a publicly available data source that allows users to conduct a search for individuals using various criteria such as birth date, death date, first name, last name, Social Security number, and place of last residence. The DMF records deaths that occurred from 1937 to the present and captures about 95% of deaths recorded by the National Death

Index (Sesso, Paffenbarger, and Lee 2000). Many researchers suggest that Social Security Administration data for older persons are superior to vital statistics records because the application processes for Social Security numbers and Medicare have strict evidentiary requirements, whereas age reporting for death certificates relies on proxy informants (Faig 2001; Kestenbaum 1992; Kestenbaum and Ferguson 2001; Rosenwaike and Stone 2003).

We established definite match when information on first and last names (spouse's last name for women) and day, month, and year of birth agreed in both DMF and family history (Sesso, Paffenbarger, and Lee 2000). In the case of disagreement in day, month, or year of birth, the validity of the match was verified on the basis of agreement between place of last residence and place of death.

The lack of a match between family history and DMF records could occur for a number of reasons: a misprint in genealogy, a missing Social Security record (particularly if the person did not use Medicare benefits), difficulty in matching a person with a common name in the absence of identical dates, and so on. The DMF covers about 90 percent of all deaths for which death certificates are issued (see Faig 2001) and about 92 to 96 percent of deaths for persons older than 65 years (Hill and Rosenwaike 2001). Further investigation of nonmatched cases using additional data sources (e.g., obituaries, state collections of death certificates) revealed that about half of nonmatched cases were related to misprints in genealogies and about 20 percent of nonmatched cases contained correct death dates (as confirmed by linkage to state death indexes), although they were not recorded in the DMF.

It should be noted that the success rate for linkage to DMF records was substantially higher for persons born after 1889, at 82 percent. This result is consistent with previous studies that reported lower quality and coverage of the DMF database for persons born before 1890 (Faig 2001).

Next, birth dates for the 534 records of persons with confirmed centenarian status born after 1889 and matched to the DMF were verified through linkage to early census data. This verification was accomplished using data from the 1900 U.S. census recorded when the person was a child (when age exaggeration is less common, compared to claims of exceptional longevity made at old age). Preference was given to the 1900 census because of its greater completeness and detail in regard to birth date verification (containing both month and year of birth) compared to the 1910 and 1920 censuses. If a person could not be found in the 1900 census, then we searched for him or her in the 1910, 1920, and 1930 censuses. The online availability of the entire indexed U.S. 1900, 1910, and 1920 censuses—as provided by Genealogy.com and Ancestry.com—and supplementary information in family histories allowed us to obtain a good linkage success rate of 91 percent. Using the census indexes, we conducted a search on the following variables: first names, last names (including Soundex, a four-character code to evaluate the similarity of pronounced words), state, county, township, birthplace, birth year (estimated on the censuses), immigration year, and relation to head-of-household. A definite match was established when information on parents and all their children (name, order, age, place of birth) were the same in both census and family history records (a one-year disagreement in ages for some children was allowed). Possible matches were established when most of these components matched but one or two variables disagreed. Children residing outside the parental household (e.g., boarding school, hospital) were identified on the basis of their name, month, and year of birth, place of birth, and proximity to parental household. For possible matches, additional attempts were made to find information in supplementary sources (state birth, death, and marriage indexes; state censuses) and verify their validity.

In the overwhelming majority of cases, comparison of family history and census information on birth dates, birthplaces, and names of siblings produced unambiguous matches.

The success of this large-scale linkage effort helped us to alleviate some initial concerns related to the possibility that parents might have named children for a deceased sibling, the lack of distinction between half- and full-siblings in some census records, the practice of placing children in other households as a result of family circumstances, and the difficulties of both name similarities and the frequent name changes that were produced by "Americanization" attempts during the early 1900s.

Data consistency checks (an initial step of centenarian age validation) revealed a surprisingly small number of obvious inconsistencies in computerized genealogies. In one case, the alleged centenarian had parents with incorrect birth dates (indicating that the parents were born later than the person himself). This case was dropped from the study. In another case, the centenarian's father was unusually old (62 years) when the centenarian was born. This situation is not impossible, however, so we retained this case for further validation. (Because this case was later confirmed through the DMF but not found in early censuses, it was ultimately not included in the final analyses.) No other records revealed obvious inconsistencies in event dates, so 990 records were passed on for further verification.

Of the 671 centenarians found in family histories, 551 (82 percent) were successfully linked to the DMF. Centenarian status for 534 of these individuals (80 percent) was confirmed—that is, we confirmed that according to their DMF death date, these individuals lived more than 100 years. For 17 persons, the death year reported in the DMF was one to two years lower than the death year reported in the genealogical record, indicating that they lived fewer than 100 years, so these individuals were excluded from later steps of the verification process. Of the 534 records with validated death dates, 485 (91 percent) were successfully linked to early U.S. census records. The success rate for linking records found in the DMF to early U.S. census records was 91 percent, which was significantly higher than rates reported by previous studies, which ranged between 39 and 75 percent (Guest 1987; Rosenwaike and Stone 2003; Rosenwaike et al. 1998). The overall number of validated records in this age validation study was also acceptable, at 72 percent. Table 1 shows the results of the linkage process between genealogical data and DMF and early census data.

Because of the possibility of linked individuals differing from the total sample in a variety of ways, we tested the data for nonmatch bias. Table 2 compares estimates for the linked sample with the nonlinked group for selected characteristics available from online family history records. To estimate the degree of nonmatch bias, we conducted two multivariate logit models predicting the genealogical data's linkage to (1) DMF and

Table 1
Results of linking genealogical records of centenarians first to the Social Security Administration Death Master File (DMF) and then to early U.S. censuses

	Centenarians born in 1890–99		
Steps of data verification	Males	Females	Both sexes
Initial number of records	160 (100%)	511 (100%)	671 (100%)
Number of records found in DMF	130 (81%)	421 (82%)	551 (82%)
Number of records found in early censuses	115 (72%)	370 (72%)	485 (72%)

Table 2
Mean values and proportions (standard deviations) across matched and nonmatched subgroups

Characteristic	Linkage to DMF		Linkage to 1900 census	
	Matched ($N = 551$)	Nonmatched ($N = 120$)	Matched ($N = 485$)	Nonmatched ($N = 49$)
Proportion of women	0.764 (0.018)	0.758 (0.039)	0.763 (0.019)	0.837 (0.053)
Age at death	101.7 (1.2)	99.6[†] (13.8)	101.3 (1.2)	101.0 (0.8)
Birth order	3.6 (2.7)	4.3 (3.2)	3.6 (2.7)	3.6 (2.8)
Sibship size	7.1 (3.5)	7.6 (3.5)	7.1 (3.4)	6.1 (3.8)
Proportion born in the Southern state	0.347 (0.020)	0.425 (0.045)	0.328 (0.021)*	0.489 (0.072)*
Proportion ever married	0.907 (0.012)***	0.700 (0.042)***	0.907 (0.013)	0.918 (0.040)
Number of children	2.5 (2.8)	2.2 (3.0)	2.6 (2.9)	2.3 (2.5)
Paternal lifespan	73.9 (14.2)	73.4 (16.4)	74.2 (14.0)	71.9 (16.8)
Maternal lifespan	73.7 (17.3)	75.2 (19.3)	73.9 (17.2)	70.4 (20.0)

Notes: $^{*}p < .05$; $^{***}p < .001$; [†]initial selection of centenarians was based on the difference between death and birth years, so some centenarians did not survive a few months to their 100-year birthday.

(2) early census records (Rosenwaike et al. 1998). There was no significant difference between individuals who were successfully linked to DMF or early census records and nonlinked individuals regarding birth year, birth order, sibship size, age at death, parental lifespan, or gender. However, individuals linked to DMF records were more likely to have been married. Although marital status was not specifically analyzed in previous studies of DMF linkages (Rosenwaike et al. 1998), prior research has noted that death certificate information for nonmarried persons is significantly less accurate than information recorded for married persons. DMF death reporting is likely to be less accurate for single individuals rather than individuals with many relatives. Previous studies of linkages between genealogical records and early census records have revealed that a birthplace in the southern states is associated with a significantly lower likelihood of being linked to census records (Rosenwaike and Stone 2003; Rosenwaike et al. 1998), which is most likely caused by lower enumeration rates for southern states in the 1900 census (Rosenwaike and Stone 2003).

Our study observed no nonlinkage bias for the majority of the analyzed characteristics. Better representation of married individuals in the DMF and persons born in non-southern states in the census records is most likely caused by better coverage of these groups in the original data sources (DMF and censuses), rather than by the properties of the linkage procedure. Compared to other linkage studies, the percentages of nonlinked records in our sample are rather low, and they are lower than nonresponse rates in many population surveys (Groves 2006).

As our next step, we extracted 198 records of centenarians born between 1890 and 1893 from our sample of validated centenarians born between 1890 and 1899 for additional reconstruction of sibship structure and sibling lifespan in order to conduct within-family analyses. To meet this aim, we reconstructed complete family histories of the selected 198 centenarians using census records, Social Security Administration data, genealogical records, and other supplementary data resources. We reconstructed all birth dates of centenarian siblings using information available in computerized genealogies and early census records. DMF verification of death dates was not feasible for validating the death dates of shorter-lived siblings (used as controls), because DMF data completeness is not very high for deaths occurring prior to the 1970s (when systematic death date reporting began; see Faig 2001). State death indexes, cemetery records, and obituaries cover longer periods of time. Taking into account that finding exact ages of death for controls is not particularly important for purposes of comparison (it is sufficient to assume that shorter-lived siblings did not survive to age 100), we relied on death date information recorded in family histories for siblings not found in external sources. This approach was used previously in the Utah Population Database study for individuals who died before 1932 (Kerber et al. 2001).

When all reported siblings ($N = 1{,}141$) in our family history reconstruction, including those who died in childhood, were taken into account, each case (centenarian) had an average of six control siblings. This average sibship size (seven siblings) in our centenarian families is higher than the average number of children in American families reported by the 1900 U.S. census—5.12 ± 0.01 (Ruggles et al. 2010). The larger sibship size in centenarian families compared to sibship size in the general population can be explained by two factors: first, genealogies are more likely to be compiled for larger families, and second, longer-lived individuals in the United States have historically been born more often in rural areas with higher fertility rates (Gavrilova and Gavrilov 2007; Preston, Hill, and Drevenstedt 1998).

As a result of family reconstitution efforts, we were able to identify death dates for 930 siblings using DMF records, state death indexes, and online genealogies. In those cases in which information on siblings' death dates was not available, we confirmed these individuals' survival to adulthood on the basis that they had spouse(s). There were 97 such cases. Birth order of all putative first-born centenarians was verified using available data from early censuses. To control for historical changes in living conditions, only data for siblings whose birth dates did not differ by more than 10 years from the centenarian's birth date were used for further statistical analyses (763 cases).

Variables and Statistical Methods

In this study, we analyzed family variables that vary within a family and are available in computerized genealogies: maternal and paternal age at reproduction and birth order. Maternal and paternal ages were available for all siblings because parents' birth years were available in both computerized genealogies and early census records. Information about birth order was taken from genealogies and verified using early censuses (census information per se could not be used for birth order identification because of possible incompleteness of sibship in census records).

Statistical analyses of within-family effects were performed using a conditional multiple logistic regression model (fixed-effect model) to investigate the relationship between an outcome of being a case (a long-lived person) and a set of prognostic factors (Breslow and Day 1993; Hosmer and Lemeshow 2001). The fixed-effects logit model can be written as

$$\Pr(y_{it} = 1 \,|x_{it}) = F(\alpha_I + X_{it}\beta)$$

where $F(z) = \exp(z)/(1 + \exp(z))$ is a cumulative logistic function; $i = 1, 2, \ldots n$ denotes the families (independent units); $t = 1, 2, \ldots T_i$ denotes children for the ith family; and x_{it} denotes the vector of within-family covariates, including maternal age and birth order. The likelihood to survive to an advance age (i.e., to be in the long-lived group) is the dependent variable. Analyses were conducted using Stata statistical software, version 11 (StataCorp 2009). The model also included the variables for birth order, paternal age, maternal age, and sex (male or female).

Results and Discussion

We found that the odds of becoming a centenarian are 1.8 times higher for first-born children compared to their later-born siblings from exactly the same family (see Table 3).

The next question that we explored concerned the role of child mortality, which was very common a century ago, when the studied centenarians were born. If first-born children were more likely to survive to adult age because of this selective child survival, this factor might make centenarians more prevalent among first-born offspring. To test this hypothesis, we re-analyzed the data and included only those siblings who survived to adulthood (with an adult defined as a person who married or lived 20 or more years). This analysis found that even for adult persons, the odds to live to 100 are almost twice as high for first-born persons than for their later-born siblings (see Table 3).

We then explored the role of paternal and maternal ages at reproduction as a potential explanation for the birth order effect using data for siblings who survived to adulthood.

Table 3

Odds ratios (95% confidence intervals) to become a centenarian as predicted by conditional logistic regression (fixed effects): Effects of parental age and birth order

Model[†]	*N*	First-born status	Born to young father (< 25 y)	Born to young mother (< 25 y)	Female sex
Model 1: All siblings	950	1.77** (1.18–2.66)			2.47*** (1.74–3.52)
Model 2: Siblings survived to adulthood[††]	797	1.95** (1.26–3.01)			2.18*** (1.52–3.14)
Model 3: Siblings survived to adulthood	797		1.76 (0.93–3.33)		2.15*** (1.50–3.09)
Model 4: Siblings survived to adulthood	797	1.88* (1.13–3.14)	1.10 (0.52–2.32)		2.18*** (1.52–3.13)
Model 5: Siblings survived to adulthood	797	1.52 (0.93–2.48)		1.67* (1.01–2.76)	2.16*** (1.50–3.11)
Model 6: Siblings survived to adulthood	797			2.02** (1.30–3.14)	2.13*** (1.49–3.07)
Model 7: Siblings survived to adulthood; both parents survived to age 50	636			2.12** (1.27–3.54)	2.23*** (1.49–3.36)
Model 8: Siblings survived to age 75	548			1.83* (1.12–2.98)	2.15*** (1.43–3.21)

Notes: *$p < .05$; **$p < .01$; ***$p < .001$; [†]blank entries indicate that variable was not included in the model; [††]adulthood is defined as survival to age 20 or being married when death date could not be identified.

We found that young (under 25 years) paternal age had no statistically significant association with odds of survival to age 100 (see Table 3, Model 3). Ultimately, a young father's age was far less important than first-born status in predicting the chances of exceptional longevity (see Table 3, Model 4). Old paternal age (over 50 years) also had no effect on survival to age 100 (data not shown here). Thus, we did not find a significant association between paternal age and longevity in this within-family study.

Finally, we included maternal age in the analysis and found that young maternal age (under 25 years) at childbirth was the most important predictor of exceptional survival, making the birth order effect statistically insignificant. These findings indicate that the beneficial effect of being first-born is driven mostly by young maternal age at a person's birth (see Table 3, Model 5). Even at age 75, being born to a young mother is still important for survival to 100 years, because the odds of exceptional survival are 1.8 times higher for first-born siblings than for later-born siblings (see Table 3, Model 8). Taking into account that birth order and parental age at reproduction may be functions of parental survival, we repeated our analyses for the offspring of parents who survived to age 50 and hence realized their reproductive potential. The results obtained for this sample (see Table 3, Model 7) are not different from the results obtained for the total sample of siblings who survived to adult age (see Table 3, Model 6). It is interesting that the survival benefits of being born to a young mother are observed only when the mother is younger than 25 years (see Figure 2).

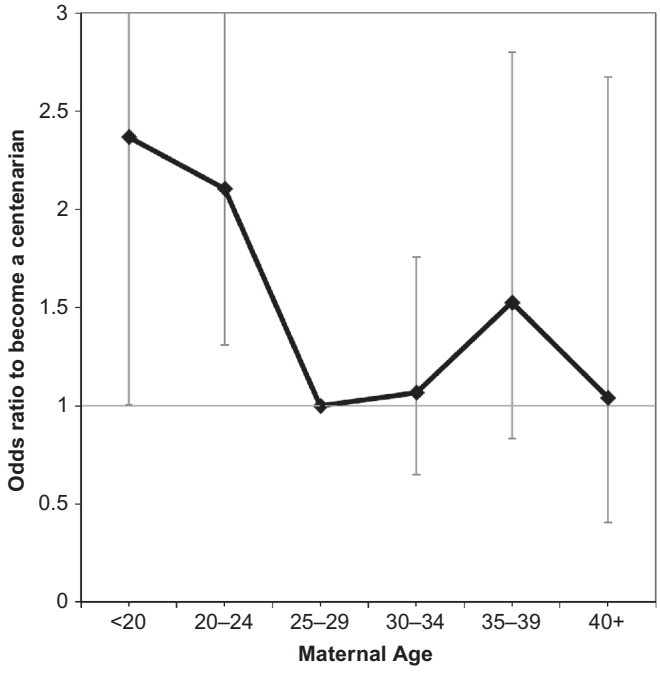

Figure 2. Odds ratio to survive to 100 years as a function of maternal age at person's birth.
Note: Data for persons born to a mother aged 25–29 years are treated as a reference category. Data are based on a logistic regression model, which includes five categories of maternal age and gender as covariates.

Thus, multivariate within-family analysis of birth order effects on human longevity revealed that a mother's young age is responsible for the apparent beneficial effects of first-born status on longevity. The within-family approach has great advantages over other methods, because it is free of the confounding caused by between-family differences. However, it remains to be seen whether the observed effect could be reproduced in further studies.

The finding of the beneficial effect of maternal age on human offspring survival to age 100 may have a biological explanation. There is empirical evidence that the quality of female eggs in humans rapidly declines with age (Bickel 2005; Pellestor, Anahory, and Hamamah 2005), and this deterioration starts early—before age 30 (Heffner 2004). Maternal age influences the biology of the mother-fetus relationship, with increasing age having a negative effect on fetal development and predisposition to severe diseases such as type I diabetes (Gloria-Bottini et al. 2005).

Experiments on laboratory mice have found that the offspring born to younger mothers live longer, with the largest effects being observed at later life (Tarin et al. 2005). Other studies have found that hormonal profiles in pregnant mice differ depending on maternal age (Wang and vom Saal 2000). This may explain why adult offspring of adolescent (35 days old) and middle-aged (9 months old) mothers have lower body weight and more delayed puberty than offspring born to young adult mothers (3 months old), and why, similarly, male offspring of this same group have smaller reproductive organs than offspring born to young adult mothers (Wang and vom Saal 2000). Wang and vom Saal (2000) have further demonstrated a transgenerational effect of maternal age in mice, with the birth weight of offspring depending on their grandmothers' age at pregnancy (Wang and vom Saal 2000). Delayed motherhood in mice has also been demonstrated to have negative effects on behavioral traits of young adult offspring (Tarin et al. 2003).

Data on the long-term effects of maternal age on human offspring are scarce. One study showed that the lifespan of children decreases with increasing maternal age (Kemkes-Grottenthaler 2004). Our earlier studies have not detected an association between maternal age and offspring mortality in historical populations of European aristocracy (Gavrilov and Gavrilova 1997; Gavrilov and Gavrilova 2000), but we believe that this might be due to some limitations in the tools used to analyze the data, such as the absence of a control for within-family variation. These earlier between-family studies also found that an older paternal age is related to a lower lifespan for adult daughters (Gavrilov and Gavrilova 2000; Gavrilov et al. 1997). Smith and colleagues also found a detrimental effect of very old paternal age (70 or more years) on the mortality of daughters using between-family analysis of data from the Utah Population Database (Smith, Mineau et al. 2009). The authors admit the difficulty of determining whether this paternal age effect is due to a higher level of germ-line mutations in the sperm of old fathers or whether it is due to problems of having been reared by old parents. This study showed a very moderate negative effect of old maternal age (35 or more years) on the survival of sons (Smith, Mineau, et al. 2009).

The fact that offspring's lifespan depends on maternal age at their birth, even in laboratory animals, indicates that some fundamental biological mechanisms may be involved. One plausible hypothesis posits that changes in genomic imprinting in oocytes of aging females may be a factor in determining lifespan (Comings and Macmurray 2001; Comings and MacMurray 2006). Another plausible biological hypothesis concerns the "telomere theory of reproductive senescence" in females (Keefe et al. 2005), which posits that eggs ovulating in older females have shorter telomeres because of late exit from the oogonial "production line" (Polani and Crolla 1991) during fetal life, with incomplete restoration of

telomere lenght by telomerase (Keefe et al. 2005). Telomeres are DNA repeats that cap and protect chromosome ends, so that longer telomeres in the eggs of younger females may be beneficial for offspring lifespan. However, in human beings, sociobehavioral mechanisms may also be involved in lifespan determination, in addition to more general biological mechanisms. For example, it is possible that maternal-age effects in human beings are mediated through the length of mothering (i.e., duration of maternal care and supervision). Children born to young mothers (under 25 years) are exposed to maternal care for a longer time on average than later-born offspring. One study has shown that maternal support is not distributed evenly within a sibship, with daughters often being given a preference (Suitor, Pillemer, and Sechrist 2006). The mothering hypothesis also implies the need for close geographical proximity between a mother and her grown child to allow long-term maternal care and influence to manifest itself through personal interactions. For example, Smith and colleagues found that "mothers who are not physically proximate to their adult offspring elevate the risks of mortality of the offspring" (Smith et al. 2005). Another possible sociobehavioral explanation of observed maternal age effects in human beings is the mediation of these effects through the size of an individual's social support group (numbers of younger siblings) that could provide care and support in later life. More research is needed to disentangle the role of biological and social mechanisms in maternal age effects on human longevity.

Study 2: Midlife Physical Characteristics and Longevity

The incorporation of physical characteristics into demographic analysis of mortality widens the scope of explanatory variables in biodemographic research on health outcomes (Crimmins and Seeman 2000). For example, early studies linked low birth weight to increased mortality from cardiovascular disease later in life (Barker 1992; Barker et al. 1993), but later research demonstrated that the relationships between birth weight, adult age adiposity, and late-life diseases are complex, with both poor growth during fetal life and infancy and rapid catch-up growth and childhood weight gain contributing to subsequent disease risk (Wells 2007). Unfortunately, detailed birth weight information is not yet available for historical populations and centenarian cohorts. However, an individual's height at young adult age seems to be a good indicator of his or her nutritional and infectious disease history, at least in historical data (Alter 2004; Alter, Neven, and Oris 2004; Elo and Preston 1992). Most studies have found a negative relationship between body height and mortality later in life (Elo and Preston 1992; Waaler 1984), including studies of Union Army veterans (Costa 1993b; Costa and Lahey 2005; Fogel and Costa 1997) and studies of modern Norwegian males (Costa 1993a; Costa 1993b; Costa and Lahey 2005; Fogel and Costa 1997). Infectious diseases (and diarrheal diseases in particular) can result in growth retardation, leading to shorter adult height. For example, conscripts from high-mortality districts of antebellum New York were shorter than those from healthier districts (Haines, Craig, and Weiss 2003).

Adult body height is affected by both environmental (early-life nutrition and exposure to infections) and genetic factors. Lauderdale and Rathouz (1999) have hypothesized that familial resemblance in height was suppressed in the past because of early environmental effects. Preston and Haines (1991) have suggested that the U.S. population at the end of the nineteenth century had relatively good nutritional status but a very high burden of infections. Thus, we may hypothesize that the low height of men born at the end of the nineteenth century is related to the infectious diseases from which they suffered during

childhood. If the hypothesis that childhood infections may cause late-life chronic diseases is correct (Finch and Crimmins 2004), we may expect centenarians at young adult ages to be taller on average than their peers who did not survive to advanced ages. According to this hypothesis, "chronic inflammatory mechanisms drive much of the influence of early-life infections on later morbidity and mortality" (Finch and Crimmins 2004), and "height is also linked to infections and the inflammatory response [in childhood]," because "if infections occur during development, substantial energy is reallocated at the expense of growth, as required by the body for immune defense reactions and for repair" (Crimmins and Finch 2006). Thus, one may expect centenarians to be taller at young adult ages than their non-centenarian peers, because they should have experienced fewer childhood infections, which are detrimental to both body growth and subsequent longevity.

An alternative view has been suggested by Samaras, who believes that rapid growth and maximum height attainment are not desirable goals (Samaras 2009; Samaras and Storms 2002). His findings suggest that men with lower height and weight live on average longer than their taller and heavier peers (Samaras and Storms 1992; Samaras, Storms, and Elrick 2002). Studies of different strains and breeds of dogs, rats, and mice have also showed that smaller animals live on average longer than their larger counterparts within a given species (Li et al. 1996; Michell 1999; Miller, Chrisp, and Atchley 2000; Patronek, Waters, and Glickman 1997; Samaras, Elrick, and Storms 2003). Some biologists believe that rapid growth may be harmful, and that somewhat delayed maturation may be beneficial for longevity and health (Rollo 2002).

Centenarians' body size during their young years remains unclear. Historical studies suggest that centenarians should be on average much taller than their peers as a result of better nutrition and avoidance of diseases early in life. However, existing biological data predict that centenarians should have been shorter than the average height. Chan, Suzuki, and Yamamoto (1999; 1997a; 1997b) have reported that Japanese centenarians are shorter than persons of average lifespan, although these studies measured centenarians at old age, when their height had already decreased. Studies of centenarian body size at young adult age would help to resolve the existing controversy regarding the relationship between body size and longevity.

In this study, we compare male centenarians at age 30 to their shorter-lived peers using data from the U.S. World War I Civilian Draft Registration Card register, available online through Ancestry.com. The uniqueness of this data source lies in its wealth of information on the physical characteristics of draft-registered men, including data on body height and build. These data allowed us to analyze the effects of physical characteristics at a young age on longevity.

Study Design and Data

In this study, we applied a matched case-control design in which shorter-lived men were matched with centenarian men by birth year, race, and county of draft registration and used as controls (see Figure 3). This approach allowed us to eliminate the effects of birth cohort, race, and place of draft registration on survival. The use of controls from the same geographical area (county) allowed us to mitigate a possible geographically related subjectivity in height and build estimation.

We developed the study sample in three stages. In the first stage, we randomly selected the records of 240 men born in 1887 and surviving to age 100 from the DMF. We used the 1887 birth cohort to avoid possible effects of birth year "heaping" (i.e., rounding up birth years to end in 5 or 0). Furthermore, men born in 1887 reached age 30 in 1917, so their

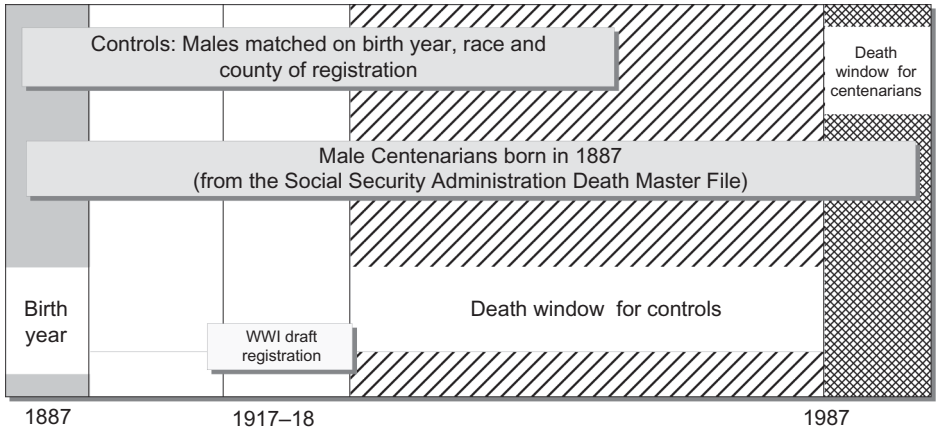

Figure 3. Matched case-control design of the study.

adult height would have been attained by the time of their draft registration. Given that the DMF covers 93 to 96 percent of the deaths of persons aged 65 years and over (Hill and Rosenwaike 2001), it was possible to apply a simple random sampling design to the data. We can consider the 1887 birth cohort to be practically extinct now, because it is highly unlikely that any man born in 1887 would live more than 120 years. Thus, we may expect the DMF to contain records on almost all American centenarians born in 1887—another advantage of selecting this birth cohort. The DMF database contains about 2,500 death records of male centenarians born in 1887, so linking all of these entries to WWI civilian draft registration cards would require a significant investment of time and effort. Therefore, we opted to use a randomly selected sample of 240 (9.6 percent of the total sample) male centenarians born in 1887. In the second stage, we linked the selected records to the WWI civilian draft registration cards. In the third stage, we matched each centenarian record to a control record randomly selected from the civilian draft registration records of persons of the same birth year, race, and county of registration.

Brief Description of WWI Draft Registration Cards

In 1917 and 1918, approximately 24 million men born between 1873 and 1900 completed draft registration cards. Men already on active duty in the military were excluded from draft registration. Registration of eligible men has been determined to be close to 100 percent, meaning that about 98 percent of adult men under age 46 living in the United States in 1917–18 completed registration cards (Banks 2000). Instructions for filling in each question on the card were posted at all registration sites, and some local newspapers printed copies of sample cards in the days prior to registration. In the vast majority of cases, volunteer staff at local offices filled in card information and then had the registrant sign his name. A more detailed description of this data source is available in Banks (2000). Table 4 presents the information listed in the draft registration cards.

The linkage process was facilitated by the availability of online indexes and actual digitized images of draft registration cards, as well as the listing of registrants' exact birth date (day, month, and year) and name both on the WWI draft cards and in the DMF records. These circumstances allowed us to obtain unambiguous matches in the majority of cases.

Table 4
Information available from WWI draft registration cards

Group	Description
Core demographic data	Age, date/place of birth, race, citizenship
Geographical data	Permanent home address
Working characteristics	Occupation, employer's name
Family characteristics	Marital status, information about dependents (including children below age 12)
Physical characteristics	Height (tall, medium, short), build (slender, medium, stout), eye color, hair color, baldness, disability

Model Specification and Statistical Methods

Variables.

Physical characteristics. Draft registration cards reported three categories of body height (tall, medium, or short) and three categories of body build (slender, medium, or stout), which were this study's main variables of interest. Additional physical characteristics reported on the cards include eye color, hair color (or baldness), and, in the case of inability to participate in military service as a result of disability, the type of disability. However, these cases turned out to be very rare in our sample.

Race. Race was used as a basis for matching and was not used in statistical analyses.

Place of birth. Place of birth reported in draft registration cards allowed us to identify whether an individual was foreign-born or native-born.

Marital status. Draft registration cards either recorded that the person was married or mentioned a wife among the person's dependents.

Number of children. Draft registration cards recorded all persons' children aged under 12 years. Given the unlikelihood of 30-year-old men having children older than 12 (i.e., becoming a father before age 18), we may suggest that draft registration cards reported almost all existing children for men in our sample.

Occupation. Draft registration cards reported registrants' current occupation. We classified occupations into five groups: farmers, white-collar workers (e.g., clerks, bankers), skilled blue-collar workers (e.g., repair mechanics, machinists), service workers (e.g., grocers, barbers, salesmen), and unskilled workers (e.g., laborers, kitchen hands).

Statistical methods. Statistical analyses were performed using a conditional multiple logistic regression model for matched case-control studies to investigate the relationship between an outcome of being a case (survival to age 100) and a set of predictor variables (Breslow and Day 1993; Hosmer and Lemeshow 2001). An important advantage of conditional logistic regression is its high statistical power (Woodward 2005), which allows researchers to detect statistically significant effects even in samples of a relatively small size.

When each matched set consists of a single control (1–1 matched study), the conditional likelihood denoted as:

$$\prod_i \left(1 + \exp(-\beta'(x_{i1} - x_{i0})\right)^{-1}$$

where x_{i1} and x_{i0} are vectors representing the prognostic factors for the case and the control, respectively, of the ith matched set (Hosmer and Lemeshow 2001). We preselected a subset of explanatory variables for possible inclusion in a multivariate model on the basis of their univariate analysis and conducted computations using Stata software, release 11 (StataCorp 2009).

Results and Discussion

The overall linkage rate for DMF-reported centenarians and draft registration card data was 72.5% (174 linked records). It should be noted that not all centenarians found in the DMF could participate in the WWI draft registration. Study of additional data sources revealed that two persons in the DMF sample were already serving in a regular army during the draft registration and seven persons had Social Security numbers that were issued after 1955, suggesting late immigration. In six cases, we found misprints in DMF records (according to their death certificates, persons were in fact born in 1987). Elimination of these noneligible cases increased the linkage success rate to 77.3 percent. Further analysis revealed a very high proportion of persons with Eastern European, Italian, and Spanish surnames among nonlinked records (41 percent), compared to persons linked to the WWI draft registration records (only 9 percent). This suggests that many individuals in the nonlinkage group could have immigrated to the United States after 1917. This suggestion was further confirmed by information on foreign-born status among draft registration controls. Table 5 describes the demographic and socioeconomic characteristics of centenarians (cases) and controls.

Note that the proportion of foreign-born individuals is similar for both cases and controls. Thus, we may conclude that the linkage rate for centenarian cases and WWI draft registration card records was not lower for foreign-born individuals than for native-born

Table 5
Characteristics of men born in 1887 and participating in the World War I civil draft registration

	Proportion (percent)		
Characteristic	Centenarians (cases; $n = 171$)	Controls ($n = 171$)	p
Race			
White	93.57	93.57	—
Black	5.26	5.26	—
Other	1.17	1.17	—
Foreign-born	20.47	22.22	.692
Married	68.42	63.74	.361
Had children	52.63	42.11	.051
Farmer by occupation	31.55	23.35	.093
Reported disability	7.02	8.77	.547
Tall height	28.07	34.50	.200
Stout build	7.02	14.62	.024

persons. Proportions of foreign-born individuals in our sample are very close to the proportions noted in official data. For example, the 1920 U.S. census reports the proportion of foreign-born individuals in age group 20–44 as 17.7 percent (U.S. Department of Commerce, 1940), which is close to our estimates. According to the same census, the proportion of black individuals in this age group was 9.8 percent. Taking into account the higher mortality of blacks compared to whites, it is reasonable to expect a decreasing proportion of blacks among centenarians, as is demonstrated by our sample (see Table 5). Comparison to official data suggests that the linkage of centenarian records to WWI draft registration cards was not subject to significant biases regarding foreign-born status or race.

Table 5 also contains a distribution of cases and controls according to individuals' body height and build. Note that the "tall" category corresponds to the top 35th percentile of the tallest men in the control population. A "stout" body build corresponds to the top 15th percentile of the heaviest men in the control population, whereas this proportion is significantly lower in the centenarian group (see Table 5). Figure 4 shows the distribution of long-lived and control groups according to individuals' height at age 30.

It is interesting to note that centenarians were not concentrated among the tallest men measured at age 30. In fact, most centenarians were of medium height, although these differences were not statistically significant. Distribution of centenarians and controls by their body build at age 30 is presented in Figure 5. Only 7 percent of the future centenarians fell into the "stout" category, compared to 15 percent of the control group. The difference in body build distribution of cases and controls was found to be significant in univariate analyses (see Table 5). Multivariate analyses using conditional logistic regression found that stout body build had a statistically significant association with lower survival rates to age 100 in all three models (see Table 6).

Thus, the study of height and build of men born in 1887 suggests that obesity at a young adult age (defined as 30 years) is detrimental to the attainment of exceptional longevity, whereas body height is a far less important predictor of exceptional longevity. The finding that stout body build predicts a much lower rate of survival to 100 is generally consistent with existing knowledge about the association between high body mass index and obesity and increased mortality (Adams et al. 2006; Flegal et al. 2005; Flegal et al. 2007). Our findings also expand this knowledge further in three ways. First, the detrimental effects of obesity may have an exceptionally long time range; that is, obesity at a young

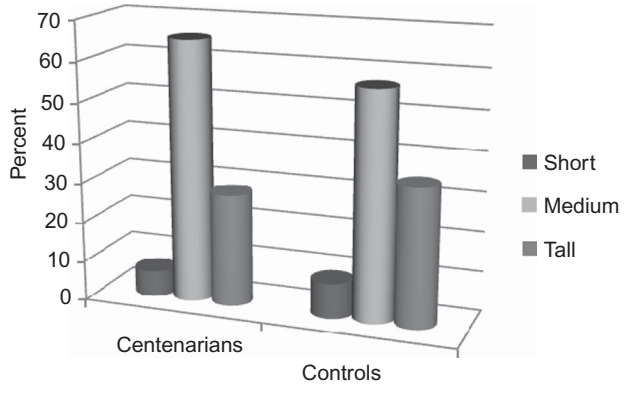

Figure 4. Body height at age 30 and survival to age 100, with distribution of cases (future centenarians) and controls by height.

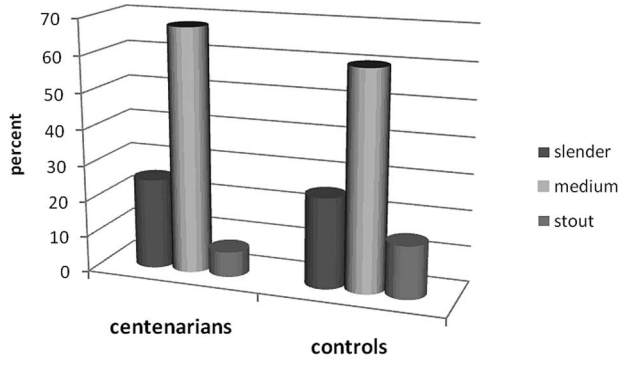

Figure 5. Body build at age 30 and survival to age 100. Distribution of cases (future centenarians) and controls by the body build category.

Table 6
Odds ratios (95% confidence intervals) of exceptional longevity (survival to age 100) for certain physical and socio-demographic characteristics of men at age 30 using multivariate conditional logistic regression

Characteristic	Model 1	Model 2	Model 3
Stout body build	Reference	Reference	Reference
Slender or medium build	2.62* (1.19–5.77)	2.63* (1.17–5.89)	2.63* (1.13–6.12)
Farmer by occupation vs. other occupation	2.00* (1.09–3.64)	2.03* (1.09–3.78)	2.20* (1.16–4.19)
Native born vs. foreign born		1.12 (0.63–1.99)	1.13 (0.63–2.05)
Married vs. nonmarried		0.76 (0.41–1.44)	0.68 (0.35–1.34)
No children		Reference	Reference
1–3 children		1.62 (0.89–2.95)	1.61 (0.87–2.98)
4+ children		2.71* (0.99–7.39)	2.59†† (0.92–7.28)
Short height			Reference
Medium or tall height			1.35 (0.80–2.29)
Blue/gray eyes			1.71† (0.99–2.95)
Light hair			0.64 (0.31–1.32)
Disability			0.68 (0.28–1.66)

Notes: $^{*}p \leq .05$; $^{\dagger}p = .052$; $^{\dagger\dagger}p = .07$.

adult age (30 years) is still predictive for decreased chances of survival to age 100. Second, the significance of body build as a predictor of exceptional longevity is much higher than the significance of all other potentially important variables, such as body height, immigration status, marital status, and occupation (with the exception that being a farmer is highly beneficial for attaining exceptional longevity). Third, contrary to expectations based on life

extension of calorically restricted animals (Fontana, Partridge, and Longo 2010), a "slender" body build does not improve chances of survival to 100 years. It should be noted that slender body build in the past could have been related to poor nutrition or infectious load, tuberculosis in particular (Alter 2004; Elo and Preston 1992), but our data do not support the hypothesis of decreased chances of exceptional longevity for slender-built individuals.

Our study also found that body height is not associated with survival to age 100, and that centenarians tend to have medium height on average. Thus, our data do not support either of our hypotheses related to adult body height and longevity described earlier in this article. It should be noted, however, that the absence of an effect of body height on longevity in our study might be related to the study's small sample size and resulting insufficient statistical power to reveal potential height-longevity effects. Also, other studies have shown that tall body height may have multidirectional associations with chronic diseases, including a positive association with cancer (Batty et al. 2009) and a negative association with heart disease (Paajanen et al. 2010). Therefore, the final effect of tall body height on longevity may be weak and more complicated than the effect of obesity on longevity. One recent study of height and late-life mortality among Finnish men showed that persons who were taller than their peers at age seven lived longer (Barker et al. 2011). However, one group of men with unexpectedly tall height at age seven had elevated mortality in middle age (Barker et al. 2011), indicating the detrimental effect of early compensatory growth on longevity.

Another finding of this study is the positive effect of a farmer occupation on survival to age 100. No other occupations studied here, including white-collar work, had a significant effect on the attainment of longevity. This result is consistent with our previous findings suggesting that children raised on farms (boys in particular) have higher chances of becoming centenarians (Gavrilova and Gavrilov 2007). Similar results have been obtained by other authors studying childhood conditions and survival to advanced ages (Hill et al. 2000; Preston, Hill, and Drevenstedt 1998; Stone 2003). Preston and colleagues (Preston, Hill, and Drevenstedt 1998) have hypothesized that the effect of a farm childhood on longevity is stronger for men than women because men raised on farms often become farmers by occupation and continue to live in this healthy environment. Our findings here are consistent with this hypothesis.

Being married by age 30 had no statistically significant effect on survival to age 100. However, the number of children that an individual had at age 30 demonstrated positive effects on chances of exceptional longevity (see Table 6). It is interesting to note that having a large number of children born by age 30 increases a man's chances of attaining exceptional longevity by a factor of 2.6–2.7 (see Table 6). This positive association seems to differ from the predictions of some evolutionary theories of aging, such as disposable soma theory. According to this theory, for both men and women, "there may be a trade-off between reproductive success and longevity, because resources invested in longevity assurance may be at [the] expense of reproduction" (Westendorp and Kirkwood 1998). However, our finding of a positive relationship between reproductive success and longevity may have reasonable explanations of both a social and a biological nature. First, the production of a large number of offspring early in a parent's life may provide necessary caregiving and material support for a parent in his older ages. Second, high fertility at a young age may be a marker of a man's overall good health. Further studies of centenarians, including studies of genealogical data, may shed more light on the mechanisms of this interesting phenomenon.

A number of data limitations need to be considered in evaluating the results related to body build and height characteristics. Although draft registration cards contain valuable

information on individual physical markers, this resource is not free of problems. The main difficulty we faced was the categorical classification (rather than metrical measurement) of height and build data, which is less precise than those measures used in specialized health surveys like the National Health and Nutrition Examination Survey. During WWI draft registration, local staff were asked to classify individual men's height and weight. The three categories provided were rather vague, and occasionally, the staff wrote in actual weight and height instead. In addition, some errors in reporting physical characteristics also occurred (Banks 2000). Nevertheless, because volunteer staff in the registration office did collect the relevant physical data when centenarians were young adults, these data are not subject to self-report and recall bias. Furthermore, our use of county-matched controls helped us avoid possible regional differences in defining "tallness" or "shortness." This study provides the first estimates of height and build for U.S. centenarians at young ages, which may be helpful in testing alternative hypotheses on early growth and longevity, such as those proposed by Costa and Lahey (2005), Fogel (2003), Miller et al. (2002), and Willcox et al. (2006). The results of this study also demonstrate the usefulness of the U.S. WWI draft registration card dataset as a new and promising source of information for finding factors associated with lower mortality and better survival.

Conclusions

In this article, we presented results of two studies demonstrating that a matched case-control design approach is a useful way to study the effects of early-life conditions on exceptional longevity. In the first study, a sample of centenarians was drawn from a large set of computerized family histories available online. Linkage of these records to official resources provided external verification of the birth dates of both centenarians and their siblings, as well as death dates of long-lived individuals. Within-family comparison of centenarians to their shorter-lived siblings allowed us to control for unobserved common conditions early in life and common genetic background. The study demonstrated that a young maternal age is the best predictor of exceptional longevity when compared to birth order or paternal age. Young maternal age has a long-lasting effect on offspring's survival to age 100 that can be observed even after age 70. This result suggests that delayed childbearing in contemporary populations may have adverse effects on the health of future generations.

In the second study, we applied a matched case-control design to analyze the effects of physical characteristics at a young adult age on survival to age 100. Using a random sample of U.S. centenarians born in the same year, we conducted a linkage of these records to WWI civil draft registration cards collected in 1917–1918, which included information on registrants' physical characteristics. The matched case-control design allowed us to control for birth cohort, race, and contextual variables on longevity. Results showed that being overweight at a young age may significantly decrease chances of living to 100. Our results also supported previous findings that a farm upbringing improves survival to advanced ages (Gavrilova and Gavrilov 2007; Preston, Hill, and Drevenstedt 1998).

The results obtained in these studies demonstrate that factors operating early in life are important components in determining exceptional longevity. While genetic factors undoubtedly play an important role in survival to advanced ages, there are also a number of other factors associated with living to age 100 and beyond. Adding new historical data resources into our arsenal of centenarian studies will bring better understanding of the role of early-life conditions in the attainment of exceptional longevity.

References

Adams, J. W., and A. B. Kasakoff. 1984. Migration and the family in colonial New England—the view from genealogies. *J Fam Hist* 9:24–43.

———. 1991. Estimates of census underenumeration based on genealogies. *Soc Sci Hist* 15:527–543.

Adams, K., A. Schatzkin, T. Harris, V. Kipnis, T. Mouw, R. Ballard-Barbash, A. Hollenbeck, and M. Leitzmann. 2006. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *N Engl J Med* 355:763–778.

Adkins, G., P. Martin, and L. W. Poon. 1996. Personality traits and states as predictors of subjective well-being in centenarians, octogenarians, and sexagenarians. *Psychol Aging* 11:408–416.

Alter, G. 2004. Height, frailty, and the standard of living: modelling the effects of diet and disease on declining mortality and increasing height. *Popul Stud (Camb)* 58:265–279.

Alter, G., M. Neven, and M. Oris. 2004. Stature in transition—a micro-level study from nineteenth-century Belgium. *Soc Sci Hist* 28:231–247.

Alter, G., and M. Oris. 2005. Childhood conditions, migration, and mortality: migrants and natives in 19th-century cities. *Soc Biol* 52:178–191.

Anderton, D. L., L. L. Bean, J. D. Willigan, and G. P. Mineau. 1984. Adoption of fertility limitation in an American frontier population: an analysis and simulation of socio-religious subgroups. *Soc Biol* 31:140–159.

Anderton, D. L., N. O. Tsuya, L. L. Bean, and G. P. Mineau. 1987. Intergenerational transmission of relative fertility and life course patterns. *Demography* 24:467–480.

Banks, R. 2000. World War I civilian draft registrations. Online database, Ancestry.com, Provo, UT.

Barker, D. J. P. 1992. *The fetal and infant origins of adult disease*. London: BMJ Books.

Barker, D. J. P., E. Kajantie, C. Osmond, K. L. Thornburg, and J. G. Eriksson. 2011. How boys grow determines how long they live. *Am J Hum Biol* 23:412–416.

Barker, D. J. P., P. D. Gluckman, K. M. Godfrey, J. E. Harding, J. A. Owens, and J. S. Robinson. 1993. Fetal nutrition and cardiovascular disease in adult life. *Lancet* 341:938–941.

Barzilai, N., G. Atzmon, C. A. Derby, J. M. Bauman, and R. B. Lipton. 2006. A genotype of exceptional longevity is associated with preservation of cognitive function. *Neurology* 67:2170–2175.

Barzilai, N., and A. R. Shuldiner. 2001. Searching for human longevity genes: the future history of gerontology in the post-genomic era. *J Gerontol A Biol Sci Med Sci* 56:M83–M87.

Batty, G. D., M. J. Shipley, D. Gunnell, R. Huxley, M. Kivimaki, M. Woodward, C. M. Y. Lee, and G. D. Smith. 2009. Height, wealth, and health: an overview with new data from three longitudinal studies. *Econ Hum Biol* 7:137–152.

Bean, L. L., G. P. Mineau, and D. L. Anderton. 1992. High-risk childbearing: fertility and infant mortality on the American frontier. *Soc Sci Hist* 16:337–363.

Bengtsson, T., and G. R. Mineau. 2009. Early-life effects on socio-economic performance and mortality in later life: a full life-course approach using contemporary and historical sources: introduction. *Soc Sci Med* 68:1561–1564.

Ben-Shlomo, Y., and D. Kuh. 2002. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol* 31:285–293.

Bickel, S. E. 2005. Aging (not so) gracefully. *Nat Genet* 37:1303–1304.

Bottini, E., G. F. Meloni, J. MacMurray, M. Ammendola, T. Meloni, and F. Gloria-Bottini. 2001. Maternal age and traits of offspring in humans. *Placenta* 22:787–789.

Breslow, N. E., and N. E. Day. 1993. *Statistical methods in cancer research, vol.1: the analysis of case-control studies*. Oxford, UK: Oxford University Press.

Chan, Y. C., M. Suzuki, and S. Yamamoto. 1997a. Dietary, anthropometric, hematological and biochemical assessment of the nutritional status of centenarians and elderly people in Okinawa, Japan. *J Am Coll Nutr* 16:229–235.

———. 1997b. Nutritional status of centenarians assessed by activity and anthropometric, hematological and biochemical characteristics. *J Nutr Sci Vitaminol* 43:73–81.

———. 1999. A comparison of anthropometry, biochemical variables and plasma amino acids among centenarians, elderly and young subjects. *J Am Coll Nutr* 18:358–365.

Comings, D. E., and J. Macmurray. 2001. Maternal age as a confounding variable in association studies. *Am J Med Genet B Neuropsychiatr Genet* 105:564–564.

———. 2006. Maternal age at the birth of the first child as an epistatic factor in polygenic disorders. *Am J Med Genet B Neuropsychiatr Genet*. 141:1–6.

Costa, D. L. 1993a. Height, wealth, and disease among the native-born in the rural, antebellum North. *Soc Sci Hist* 17:355–383.

———. 1993b. Height, weight, wartime stress, and older age mortality: evidence from the Union Army records. *Explor Econ Hist* 30:424–449.

———. 2003. Understanding mid-life and older age mortality declines: evidence from Union Army veterans. *J Econometrics* 112:175–192.

Costa, D. L., and J. Lahey. 2005. Becoming oldest old: evidence from historical U.S. data. *Genus* 61:125–161.

Crimmins, E., and C. Finch. 2006. Infection, inflammation, height, and longevity. *Proc Natl Acad Sci U S A* 103:498–503.

Crimmins, E. M., and T. Seeman. 2000. Integrating biology into demographic research on health and aging (with a focus on the MacArthur Study of Successful Aging). In *Cells and surveys*, eds. C. E. Finch, J. W. Vaupel, and K. Kinsella, 9–41. Washington, DC: National Academies Press.

Crow, J. F. 1997. The high spontaneous mutation rate: is it a health risk? *Proc Natl Acad Sci U S A* 94:8380–8386.

———. 2000. The origins patterns and implications of human spontaneous mutation. *Nat Rev Genet* 1:40–47.

Elo, I. T., and S. H. Preston. 1992. Effects of early-life condition on adult mortality: a review. *Popul Index* 58:186–222.

Elo, I. T., S. H. Preston, I. Rosenwaike, M. Hill, and T. P. Cheney. 1996. Consistency of age reporting on death certificates and social security records among elderly African Americans. *Soc Sci Res* 25:292–307.

Faig, K. 2001. Reported deaths of centenarians and near-centenarians in the U.S. Social Security Administration's death master file. Paper presented at the Society of Actuaries "Living to 100 and Beyond" international symposium, Florida, January.

Finch, C. E., and E. M. Crimmins. 2004. Inflammatory exposure and historical changes in human life-spans. *Science* 305:1736–1739.

Flegal, K., B. Graubard, D. Williamson, and M. Gail. 2005. Excess deaths associated with underweight, overweight, and obesity. *JAMA* 293:1861–1867.

———. 2007. Impact of smoking and preexisting illness on estimates of the fractions of deaths associated with underweight, overweight, and obesity in the US population. *Am J Epidemiol* 166:975–982.

Fogel, R.W. 2003. Secular trends in physiological capital: implications for equity in health care. *Perspect Biol Med* 46:S24–S38.

Fogel, R.W., and D. L. Costa. 1997. A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. *Demography* 34:49–66.

Fontana, L., L. Partridge, and V. D. Longo. 2010. Extending healthy life span: from yeast to humans. *Science* 328:321–326.

Gavrilov, L. A., and N. S. Gavrilova. 1991. *The biology of life span: a quantitative approach*. New York: Harwood Academic Publisher.

———. 1997. Parental age at conception and offspring longevity. *Rev. Clin Gerontol* 7:5–12.

———. 2000. Human longevity and parental age at conception. In *Sex and longevity: sexuality, gender, reproduction, parenthood*, eds. J.-M. Robine, T. B. L. Kirkwood, and M. Allard, 7–31. Berlin and Heidelberg: Springer-Verlag.

———. 2001. Biodemographic study of familial determinants of human longevity. *Popul* 13:197–222.

———. 2003. Early-life factors modulating lifespan. In *Modulating aging and longevity*, ed. S. I. S. Rattan, 27–50. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Gavrilov, L. A., N. S. Gavrilova, V. N. Kroutko, G. N. Evdokushkina, V. G. Semyonova, A. L. Gavrilova, E. V. Lapshin, N. N. Evdokushkina, and Y. E. Kushnareva. 1997. Mutation load and human longevity. *Mutat Res* 377:61–62.

Gavrilov, L. A., N. S. Gavrilova, S. J. Olshansky, and B. A. Carnes. 2002. Genealogical data and the biodemography of human longevity. *Soc Biol* 49:160–173.

Gavrilova, N. S., and L. A. Gavrilov. 1999. Data resources for biodemographic studies on familial clustering of human longevity. *Demogr Res* 1:1–48.

———. 2007. Search for predictors of exceptional human longevity: using computerized genealogies and Internet resources for human longevity studies. *N Am Actuar J* 11:49–67.

Gloria-Bottini, F., E. Cosmi, M. Nicotra, E. V. Cosmi, and E. Bottini. 2005. Is delayed childbearing changing gene frequencies in western populations? *Hum Biol* 77:433–441.

Groves, R. M. 2006. Nonresponse rates and nonresponse bias in household surveys. *Public Opin Q* 70:646–675.

Guest, M. 1987. Notes from the National Panel Study: linkage and migration in the late nineteenth century. *Hist Methods* 20:63–77.

Hadley, E. C., W. K. Rossi, S. Albert, J. Bailey-Wilson, J. Baron, R. Cawthon, J. C. Christian, et al. 2000. Genetic epidemiologic studies on age-specified traits. *Am J Epidemiol* 152:1003–1008.

Hagberg, B., B. B. Alfredson, L. W. Poon, and A. Homma. 2001. Cognitive functioning in centenarians: a coordinated analysis of results from three countries. *J Gerontol B Psychol Sci Soc Sci* 56:P141–P151.

Haines, M. R., L. A. Craig, and T. Weiss. 2003. The short and the dead: nutrition, mortality, and the "antebellum puzzle" in the United States. *J Econ Hist* 63:382–413.

Hayward, M. D., and B. K. Gorman. 2004. The long arm of childhood: the influence of early-life social conditions on men's mortality. *Demography* 41:87–107.

Heffner, L. J. 2004. Advanced maternal age: how old is too old? *New Engl J Med* 351:1927–1929.

Hill, M. E., and I. Rosenwaike. 2001. The Social Security Administration's death master file: the completeness of death reporting at older ages. *Soc Secur Bull* 64:45–51.

Hill, M. E., S. H. Preston, I. Rosenwaike, and J. F. Dunagan. 2000. Childhood conditions predicting survival to advanced age among white Americans. Paper presented at the annual meeting of the Population Association of America, Los Angeles.

Hosmer, D. W., and S. Lemeshow. 2001. *Applied logistic regression*. New York: Wiley & Sons.

Hubbard, R. E., M. K. Andrew, and K. Rockwood. 2009. Effect of parental age at birth on the accumulation of deficits, frailty and survival in older adults. *Age Ageing* 38:380–385.

Jang, Y. R., L. W. Poon, and P. Martin. 2004. Individual differences in the effects of disease and disability on depressive symptoms: the role of age and subjective health. *Int J Aging Hum Dev* 59:125–137.

Jeune, B., and J. Vaupel. 1999. *Validation of exceptional longevity*. Odense, Denmark: Odense University Press.

Kalmijn, M., and G. Kraaykamp. 2005. Late or later? A sibling analysis of the effect of maternal age on children's schooling. *Soc Sci Res* 34:634–650.

Kasakoff, A. B., and J. W. Adams. 2000. The effects of migration, place, and occupation on adult mortality in the American North, 1740–1880. *Hist Methods* 33:115–130.

Keefe, D. L., S. Franco, L. Liu, J. Trimarchi, B. Cao, S. Weitzen, S. Agarwal, and M. A. Blasco. 2005. Telomere length predicts embryo fragmentation after in vitro fertilization in women: toward a telomere theory of reproductive aging in women. *Am J Obstet Gynecol* 192:1256–1260.

Kemkes-Grottenthaler, A. 2004. Parental effects on offspring longevity: evidence from 17th to 19th century reproductive histories. *Ann Hum Biol* 31:139–158.

Kerber, R. A., E. O'Brien, K. R. Smith, and R. M. Cawthon. 2001. Familial excess longevity in Utah genealogies. *J Gerontol A Biol Sci Med Sci* 56:B130–B139.

Kestenbaum, B. 1992. A description of the extreme aged population based on improved Medicare enrollment data. *Demography* 29:565–580.

Kestenbaum, B., and B. R. Ferguson. 2001. Mortality of the extreme aged in the United States in the 1990s, based on improved Medicare data. Paper presented at the Society of Actuaries "Living to 100 and Beyond" international symposium, Orlando, Florida.

Kuh, D., and B. Ben-Shlomo. 1997. *A life course approach to chronic disease epidemiology*. Oxford, UK: Oxford University Press.

Lauderdale, D. S., and P. J. Rathouz. 1999. Evidence of environmental suppression of familial resemblance: height among US Civil War brothers. *Ann Hum Biol* 26:413–426.

Li, Y., B. Deeb, W. Pendergrass, and N. Wolf. 1996. Cellular proliferative capacity and life span in small and large dogs. *J Gerontol A Biol Sci Med Sci* 51:B403–B408.

Lummaa, V., and T. Clutton-Brock. 2002. Early development, survival and reproduction in humans. *Trends Ecol Evol* 17:141–147.

Martin, P., G. da Rosa, and I. C. Siegler. 2006. Personality and longevity: findings from the Georgia Centenarian Study. *Age* 28:343–352.

Michell, A. R. 1999. Longevity of British breeds of dog and its relationships with sex, size, cardiovascular variables and disease. *Vet Rec* 145:625–629.

Miller, R. A., C. Chrisp, and W. Atchley. 2000. Differential longevity in mouse stocks selected for early life growth trajectory. *J Gerontol A Biol Sci Med Sci* 55:B455–B461.

Miller, R. A., J. M. Harper, A. Galecki, and D. T. Burke. 2002. Big mice die young: early life body weight predicts longevity in genetically heterogeneous mice. *Aging Cell* 1:22–29.

Paajanen, T. A., N. K. J. Oksala, P. Kuukasjarvi, and P. J. Karhunen. 2010. Short stature is associated with coronary heart disease: a systematic review of the literature and a meta-analysis. *Eur Heart J* 31:1802–1809.

Patronek, G. J., D. J. Waters, and L. T. Glickman. 1997. Comparative longevity of pet dogs and humans: implications for gerontology research. *J Gerontol A Biol Sci Med Sci* 52:B171–B178.

Pellestor, F., T. Anahory, and S. Hamamah. 2005. Effect of maternal age on the frequency of cytogenetic abnormalities in human oocytes. *Cytogenet Genome Res* 111:206–212.

Polani, P. E., and J. A. Crolla. 1991. A test of the production line hypothesis of mammalian oogenesis. *Hum Genet* 88:64–70.

Preston, S. H., I. T. Elo, I. Rosenwaike, and M. Hill. 1996. African-American mortality at older ages: results of a matching study. *Demography* 33:193–209.

Preston, S. H., and M. R. Haines. 1991. *Fatal years: child mortality in late nineteenth-century America*. Princeton, NJ: Princeton University Press.

Preston, S. H., M. E. Hill, and G. L. Drevenstedt. 1998. Childhood conditions that predict survival to advanced ages among African-Americans. *Soc Sci Med* 47:1231–1246.

Puca, A. A., M. J. Daly, S. J. Brewster, T.C. Matise, J. Barrett, M. Shea-Drinkwater, S. Kang, et al. 2001. A genome-wide scan for linkage to human exceptional longevity identifies a locus on chromosome 4. *Proc Natl Acad Sci U S A* 98:10505–10508.

Robine, J. M., A. Cournil, N. Henon, and M. Allard. 2003. Have centenarians had younger parents than the others? *Exp Gerontol* 38:361–365.

Rollo, C. D. 2002. Growth negatively impacts the life span of mammals. *Evol Dev* 4:55–61.

Rosenwaike, I., and M. E. Hill. 1996. The accuracy of age reporting among elderly African Americans: evidence of a birth registration effect. *Res Aging* 18:310–324.

Rosenwaike, I., M. E. Hill, S. H. Preston, and I. T. Elo. 1998. Linking death certificates to early census records: the African American matched records sample (American genealogy). *Hist Methods* 31:65–74.

Rosenwaike, I., and L. F. Stone. 2003. Verification of the ages of supercentenarians in the United States: results of a matching study. *Demography* 40:727–739.

Ruggles, S., J. T. Alexander, K. Genadek, R. Goeken, M. B. Shroeder, and M. Sobek. 2010. Integrated Public Use Microdata Series (IPUMS), Version 5.0, University of Minnesota, Minneapolis, MN.

Samaras, T. T. 2009. Should we be concerned over increasing body height and weight? *Exp Gerontol* 44:83–92.

Samaras, T. T., H. Elrick, and L. H. Storms. 2003. Is height related to longevity? *Life Sci* 72:1781–1802.

Samaras, T. T., and L. H. Storms. 1992. Impact of height and weight on life-span. *Bull. World Health Organ* 70:259–267.

———. 2002. Secular growth and its harmful ramifications. *Med Hypotheses* 58:93–112.

Samaras, T. T., L. H. Storms, and H. Elrick. 2002. Longevity, mortality and body weight. *Ageing Res Rev* 1:673–691.

Sesso, H. D., R. S. Paffenbarger, and I. M. Lee. 2000. Comparison of national death index and World Wide Web death searches. *Am J Epidemiol* 152:107–111.

Shrestha, L. B., and I. Rosenwaike. 1996. Can data from the decennial census measure trends in mobility limitation among the aged? *Gerontologist* 36:106–109.

Smith, K. R., A. Gagnon, R. M. Cawthon, G. P. Mineau, R. Mazan, and B. Desjardins. 2009. Familial aggregation of survival and late female reproduction. *J Gerontol A Biol Sci Med Sci* 64:740–744.

Smith, K. R., G. Garibotti, A. Fraser, and G. P. Mineau. 2005. Adult mortality and geographic proximity of parents in Utah in the 19th and 20th centuries. Paper presented at the IUSSP Symposium on Kinship and Demographic Behavior, Salt Lake City, Utah.

Smith, K. R., G. R. Mineau, G. Garibotti, and R. Kerber. 2009. Effects of childhood and middle-adulthood family conditions on later-life mortality: evidence from the Utah Population Database, 1850–2002. *Soc Sci Med* 68:1649–1658.

StataCorp. 2009. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP.

Stone, L. 2003. Early-life conditions and survival to age 110 in the U.S. Paper presented at the "Early Life Conditions and Longevity: Reconstructive Lives from Cradle to Grave" workshop, Geneva.

Suitor, J. J., K. Pillemer, and J. Sechrist. 2006. Within-family differences in mothers' support to adult children. *J Gerontol B Psychol Sci Soc Sci* 61:S10–S17.

Tarin, J. J., V. Gomez-Piquer, C. Manzanedo, J. Minarro, C. Hermenegildo, and A. Cano. 2003. Long-term effects of delayed motherhood in mice on postnatal development and behavioural traits of offspring. *Hum Reprod* 18:1580–1587.

Tarin, J. J., V. Gomez-Piquer, F. Rausell, S. Navarro, C. Hermenegildo, and A. Cano. 2005. Delayed motherhood decreases life expectancy of mouse offspring. *Biol Reprod* 72:1336–1343.

U.S. Department of Commerce. 1940. Statistical Abstract of the United States 1939. Washington, DC: United States Government Printing Office.

Waaler, H. T. 1984. Height, weight and mortality: the Norwegian experience. *Acta Medica Scandinavica* 215:1–56.

Wang, M. H., and F. S. vom Saal. 2000. Maternal age and traits in offspring: the timing of a mouse's first litter influences the development of her pups. *Nature* 407:469–470.

Wells, J. C. K. 2007. The programming effects of early growth. *Early Hum Dev* 83:743–748.

Westendorp, R. G. J., and T. B. L. Kirkwood. 1998. Human longevity at the cost of reproductive success. *Nature* 396:743–746.

Willcox, D. C., B. J. Willcox, H. Todoriki, J. D. Curb, and M. Suzuki. 2006. Caloric restriction and human longevity: what can we learn from the Okinawans? *Biogerontology* 7:173–177.

Woodward, M. 2005. *Epidemiology: study design and data analysis*. Boca Raton, FL: Chapman & Hall/CRC.

Young, R. D., B. Desjardins, K. McLaughlin, M. Poulain, and T. T. Perls. 2010. Typologies of extreme longevity myths. *Curr Gerontol Geriatr Res*.